

Application No. 09/928,619

REMARKS

The FINAL Office Action of June 1, 2004 has been carefully considered. Reconsideration of this application, as amended, is respectfully requested. Claims 1-13, 15-22 and 25-29 are pending in this application. Of these, claims 1, 11, 22, and 28 are independent claims. An Amendment faxed March 15, 2004 amended claims 1, 5-7, 20-22 and 25; added new claims 27, 28, and 29; and canceled claims 14, 23 and 24. This Amendment amends claim 29 to add an omitted colon.

1. Response to Rejection Under 35 USC 102

The Office Action in section 5, starting on page 3, finally rejects claims 1-13, 15-22 and 25-29 under U.S.C. §102(e) as being anticipated by Shetty et al., US 2003/0046253 A1 (hereinafter referred to as "Shetty").

Applicant traverses the final rejection of claims 1-13, 15-22 and 25-29 under 35 U.S.C. §102(e) as being anticipated by Shetty and submits that Schetty does not anticipate each and every element of Applicant's invention for the reasons set forth below.

1.A. Summary of Applicant's Invention

Applicant's invention concerns a text categorizer that is adapted to classify a text object into one or more classes. The text categorizer includes a pre-processing module, a knowledge base, and an approximate reasoning module. The pre-processing module performs feature extraction, feature reduction, and fuzzy set generation to represent an unlabelled text object in terms of one or more fuzzy sets. The approximate reasoning module uses a measured degree of match between the one or more fuzzy set and categories represented by fuzzy rules in the knowledge base to assign labels to text objects for those categories that satisfy a selected decision making rule.

1.B. Summary of Shetty Reference

Shetty discloses a neuro/fuzzy hybrid approach to clustering data. In particular, Shetty discloses a technique for clustering data in which: samples of a predetermined window length are received. The sample data is checked for uncertainty and/or robustness. The data is then clustered based on the outcome of the checking. (see paragraphs 0009 – 0011 of Shetty). Shetty discloses that the

Application No. 09/928,619

data may be clustered using either a fuzzy logic approach (see paragraphs 0048 through 0078) or an unsupervised learning approach (see paragraphs 0079 – 0097).

More specifically, Shetty discloses in paragraph 0102 a summary of how its fuzzy logic approach is applied to cluster input data.

In some embodiments, analyzer 560 clusters the read data using a fuzzy logic approach by forming a compatibility relationship matrix including the read data using a distance function such that the value obtained using the distance function is between 0 and 1. Analyzer 560 then finds a transitive closure of the formed compatibility relations matrix, and computes a threshold value based on the read data to set granularity for the clustering, and further forms a binary tree including hierarchical clusters from the found transitive closure, and builds clusters based on the formed binary tree and outputs the clustered data 580. The fuzzy logic approach is discussed in more detail with reference to FIG. 2.

1.C Shetty Fails To Disclose or Suggest constructing a document class fuzzy set as recited in claims 1 and 22

The Office Action on page 3 asserts independent claim 1 is anticipated by the paragraphs 0006, 0025, 0048, 0050, and 0102 of Shetty. In particular, paragraph 0102, which is reproduced above, summarizes Shetty's "fuzzy logic approach" to *clustering data*. Furthermore in asserting claim 1 is anticipated by Shetty, the Office Action sets forth "Examiner's Notes" on page 3 of the Office Action and "Examiner's Response" on page 12, which Applicant addresses below as they apply to Applicant's claimed invention recited in independent claim 1.

Unlike the fuzzy logic approach to clustering data disclosed in Shetty, Applicant's independent claim 1 recites a method for *classifying* a text object by constructing a *document class fuzzy set* with a plurality of features extracted from the text object. Each extracted feature has a degree of membership in the document class fuzzy set and a plurality of class fuzzy sets of a knowledge base. A degree of match is measured between each of the plurality of class fuzzy sets and the document class fuzzy set. The measured degree of match is then used to assign the text object a label that satisfies a selected decision making rule.

(a) Applicant's construction of a document fuzzy set is not synonymous with Shetty's fuzzy logic approach to clustering read data (See Examiner's Note in Office Action, on page 3, lines 7-8, which refers to paragraph 0102 of Shetty):

Application No. 09/928,619

Applicant respectfully disagrees with the Office Action's assertion that Shetty's fuzzy logic approach to clustering read data is synonymous with constructing a document fuzzy set as claimed by Applicant. Shetty in paragraph 0102 clearly states that clusters are developed from read data using the following series of steps (1) forming a compatibility matrix *including the read data*, (2) finding a transitive closure of the matrix, (3) computing a threshold to set a granularity for clustering, (4) forming a binary tree from the transitive closure, and (5) building clusters based on the binary tree.

Applicant's claims do not concern the clustering of input data as recited by Shetty, instead Applicant's invention recited in claim 1 concerns classification which involves the construction of a document fuzzy set with a *plurality of features extracted from a text object* by: calculating a frequency of occurrence for each feature in the set of features in the text object; normalizing the frequency of occurrence of each feature in the set of features; and transforming the normalized frequency of occurrence of each feature in the set of features to define the document class fuzzy set.

Thus, Applicant submits that Shetty neither discloses nor suggests constructing a document class fuzzy set for a text object by transforming a normalized frequency of occurrence of each feature in a set of features extracted from the text object. Instead the method disclosed by Shetty discloses clustering data using a binary tree formed by forming a compatibility relationship matrix including read data and from which a transitive closure of the matrix is found.

Furthermore, Shetty neither discloses nor suggests using a document class fuzzy set to categorize a document as claimed by Applicant. Instead, Shetty concerns clustering data into groups, as opposed to, assigning data to a predefined set of groups. This is shown explicitly in the "crab example" disclosed in Shetty's paragraphs 0075-0078, which made up of "200 input values" (see Shetty paragraph 0022). In the crab example, the read data or input values define the clusters identified at paragraphs 0075-0078, which read data sums to $200 = 68 + 64 + 43 + 25$. If categorization were performed, the read data would be associated with existing groups.

Application No. 09/928,619

(b) Shetty's set of clusters fail to represent Applicant's knowledge base (See Examiner's Note in Office Action, on page 3, lines 11, emphasis added):

Applicant respectfully submits that a set of clusters of input values disclosed by Shetty does not represent "a plurality of sets of class fuzzy sets of a knowledge base" as claimed by Applicant. The clusters of input values identified using the method described in Shetty involves grouping input values as noted above. In contrast, Applicant makes use of a plurality of class fuzzy sets of a knowledge base to measure the degree of match with the document class fuzzy set. Also as set forth above, class fuzzy sets are not synonymous with clusters.

Applicant thus respectfully submits that the Office Action fails to identify where Shetty discloses or suggests the use of a plurality of class fuzzy sets for measuring a degree of match between a plurality of class fuzzy sets and a document fuzzy set, where the degree of match is thereafter used to assign a class label to the document fuzzy set, and where the document fuzzy set is constructed by transforming a normalized frequency of occurrence of features extracted from a text object, as claimed and disclosed by Applicant.

(c) Shetty's cluster is not synonymous with Applicant's text object label (See Examiner's Note in Office Action, on page 3, lines 15-16):

Applicant submits that associating labels with clusters is not disclosed or suggested by Shetty, as Shetty involves clustering input values to form groups, as discussed above. In contrast, Applicant's invention recited in claim 1 concerns the categorization of a text object where a measured degree of match between a plurality of class fuzzy sets and the document class fuzzy set is used to assign a label to the text object. Accordingly, Applicant's claimed invention does not concern the clustering of data into different groups as disclosed by Shetty.

(d) Membership in Shetty's clusters does not disclose or suggest Applicant's construction of a document fuzzy set (See Examiner's Note in Office Action, on page 3, lines 18-19):

Unlike Shetty's fuzzy logic approach for clustering data, Applicant's invention as recited in claim 1 includes constructing a document class fuzzy set. Specifically, Applicant's claim 1 recites that a document class fuzzy set is constructed by: calculating a frequency of occurrence for each feature in the set of features in the

Application No. 09/928,619

text object; normalizing the frequency of occurrence of each feature in the set of features; and transforming the normalized frequency of occurrence of each feature in the set of features to define the document class fuzzy set. However, as noted above in the discussion of Shetty's crab example, clusters do not identify how frequently input data occurs, they identify groups of input data.

(e) Shetty's normalizes read data based on attributes and fails to disclose or suggest normalizing frequency of occurrence of features of text objects (See Examiner's response on page 12 of the Office Action):

Turning now to "Examiner's response" set forth at the bottom of page 12, Applicant respectfully submits that claim 1 recites that features of a text object (not the text object itself) are used to construct a document class fuzzy set. Shetty discloses in paragraph 0024 that "read text data [] is transformed into unique numerical representations". In addition, Shetty discloses in paragraph 0025 that when such "read data is normalized, it "is based on attributes" of the read data. In contrast, Applicant's claim 1 recites that features are extracted from a text object. In addition, Applicant's claim 1 recites that *features extracted from a text object* are normalized, and not that a normalization is carried out that is *based on features* of a text object.

Accordingly for the reasons set forth above, Applicant respectfully submits that independent claim 1 is patentably distinguishable over Shetty. In addition, it should be noted that independent claim 22 contain the same or very similar limitations to those discussed above with respect to claim 1, and therefore the argument presented above with regard to claim 1 applies equally to independent claim 22. Insofar as claims 2-10 and 25-27 are concerned, these claims depend from one of now presumably allowable independent claims 1 and 22 and are also believed to be in allowable condition.

1.D Shetty Fails to Disclose or Suggest representing each term in a document as a granule feature fuzzy set recited in claims 11 and 28

The Office Action on pages 6-7 asserts independent claim 11 is anticipated by the paragraphs 0005, 0006, 0073, 0094, 0095, and 0102 of Shetty. In particular, paragraph 0102, which is reproduced above, summarizes Shetty's "fuzzy logic

Application No. 09/928,619

approach" to *clustering data*. Furthermore in asserting claim 11 is anticipated by Shetty, the Office Action sets forth "Examiner's Notes" on pages 6-7 and "Examiner's Response" on page 13 of the Office Action, which Applicant addresses below as they apply to Applicant's claimed invention recited in independent claim 11.

Unlike Shetty, Applicant's claim 11 recites a method for classifying a text object by computing a degree of match between each of a plurality of class *granule* feature fuzzy sets and a document *granule* feature fuzzy set to provide a degree of match for each ones of the granule features of the text object, and then aggregating each degree of match of the ones of the granule features to define an overall degree of mach for each feature of the text object.

Furthermore, Applicant submits that Shetty's clusters determined using fuzzy logic fail to disclose or suggest classification using granule features as claimed by Applicant (See Examiner's Note in Office Action, on page 6, lines 11-12, and Examiner's response on page 13 of the Office Action).

The "granularity of clustering" referred to by Shetty in paragraph 0073 concerns setting a threshold value to set, for example, cluster size. In contrast, Applicant's invention as recited in independent claims 11 and 28, concerns the extraction of "granule features" (e.g., words) from a text object, where each granule feature is represented by a plurality of fuzzy sets and associated labels (see for example Figure 11, and description in section 2.A.1.c "Granule Fuzzy Set Feature Extraction" starting on page 13, line 9 of Applicant's specification).

Moreover, Shetty's reference to "granularity of clustering" in paragraph 0073 fails to disclose or suggest as claimed by Applicant in independent claim 11: constructing a document granule feature fuzzy set; aggregating each degree of match of a plurality of ones of the granule features to define an overall degree of match for each feature; and using the overall degree of match to assign the text object a class label. That is, Applicant's use of granule feature fuzzy sets is not to "set the granularity for [] clustering" as Shetty describes in paragraph 0073 as Applicant's claimed invention does not concern clustering but instead classification. Shetty thus fails to discloses or suggests representing features of input data with a plurality of granule feature fuzzy sets as claimed by Applicant.

Application No. 09/928,619

Accordingly for the reasons set forth above, Applicant respectfully submits that independent claim 11 is patentably distinguishable over Shetty. In addition, it should be noted that independent claim 28 contain the same or very similar limitations to those discussed above with respect to claim 11, and therefore the argument presented above with regard to claim 11 applies equally to independent claim 28. Insofar as claims 12-13, 15-21, and 29 are concerned, these claims depend from one of now presumably allowable independent claims 11 and 28 and are also believed to be in allowable condition.

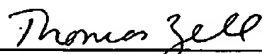
2. Fee Authorization And Extension Of Time

No additional fee is believed to be required for this amendment or response, however, the undersigned Xerox Corporation attorney hereby authorizes the charging of any necessary fees, other than the issue fee, to Xerox Corporation Deposit Account No. 24-0025. This also constitutes a request for any needed extension of time and authorization to charge all fees therefor to Xerox Corporation Deposit Account No. 24-0025.

3. Conclusion

In view of the foregoing remarks, reconsideration of this application and allowance thereof are earnestly solicited. In the event the Examiner considers a personal contact advantageous to the disposition of this case, the Examiner is hereby requested to call Attorney for Applicant(s), Thomas Zell.

Respectfully submitted,



Thomas Zell
Attorney for Applicant(s)
Registration No. 37,481
Telephone: 650-812-4282
Date: June 25, 2004